

SECURITY ANALYSIS ON “AN AUTHENTICATION CODE AGAINST POLLUTION ATTACKS IN NETWORK CODING”

JUN ZHANG, XINRAN LI AND FANG-WEI FU

ABSTRACT. We analyze the security of the authentication code against pollution attacks in network coding given by Oggier and Fathi [1] and show one way to remove one very strong condition they required. Actually, we find a way to attack their authentication scheme. In their scheme, they considered that if some malicious nodes in the network collude to make pollution in the network flow or make substitution attacks to other nodes, they thought these malicious nodes must solve a system of linear equations to recover the secret parameters. Then they concluded that their scheme is an unconditional secure scheme. Actually, note that the authentication tag in the scheme of Oggier and Fathi is nearly linear on the messages, so it is very easy for any malicious node to make pollution attack in the network flow, replacing the vector of any incoming edge by linear combination of his incoming vectors whose coefficients have sum 1. And if the coalition of malicious nodes can carry out decoding of the network coding, they can easily make substitution attack to any other node even if they do not know any information of the private key of the node. Moreover, even if their scheme can work fruitfully, the condition in their scheme $H \leq M$ in a network can be removed, where H is the sum of numbers of the incoming edges at adversaries. Under the condition $H \leq M$, H may be large, so we need large parameter M which increases the cost of computation a lot. On the other hand, the parameter M can not be very large as it can not exceed the length of original messages.

1. INTRODUCTION

Network coding is a novel technique to achieve the maximum multicast throughput, which was introduced by Ahlswede *et al.* [2]. It allows the intermediate node to generate output data by mixing its received data. In 2003, Li *et al.* [3] further showed that linear network coding is sufficient to achieve the optimal throughput in multicast networks. Subsequently, Ho *et al.* [4] introduced the concept of random linear network coding, and proved that it achieves the maximum throughput of multicast network with high probability. Network coding is efficiently applicable to numerous forms of network communications, such as Internet TV, wireless networks, content distribution networks and P2P networks. Due to these advantages, network coding attracts many researchers and has developed very quickly.

However, networks using network coding impose security problems that traditional networks do not face. A particularly important problem is the pollution attack. If some nodes in the network are malicious and inject corrupted packets into the information flow, then the honest intermediate node mix invalid packet with other packets. According to the rule of network coding, the corrupted outgoing packets quickly pollute the whole network and cause all the messages to be decoded wrongly in the destination.

This research is supported by the National Key Basic Research Program of China (Grant No. 2013CB834204), and the National Natural Science Foundation of China (Nos. 61171082, 10990011 and 60872025). The author Jun Zhang is also supported by the Chinese Scholarship Council under the State Scholarship Fund during visiting University of California, Irvine.

Recently several related works are proposed to address the pollution attack, such as homomorphic hashing, digital signature and message authentication code (MAC). Krohn *et al.* [5] (see also [6]) used homomorphic hashing function to prevent pollution attacks. Yu *et al.* [7] proposed a homomorphic signature scheme based on discrete logarithm and RSA, which however was showed insecurely by Yun *et al.* [8]. Charles *et al.* [9] gave a signature scheme based on Weil pairing over elliptic curves and provided authentication of the data in addition to detecting pollution attacks. Zhao *et al.* [10] designed a signature scheme that view all blocks of the file as vectors and make use of the fact that all valid vectors transmitted in the network should belong to the subspace spanned by the original set of vectors from the file. Boneh *et al.* [11] proposed two signature schemes that can be used in conjunction with network coding to prevent malicious modification of messages, and they showed that their constructions had a lower signature length compared with related prior work. Boneh *et al.* [12] constructed a linearly homomorphic signature scheme that authenticates vectors with coordinates in the binary field \mathbb{F}_2 . It is the first such scheme based on the hard problem of finding short vectors in integer lattices. Agrawal and Boneh [13] designed a homomorphic MAC system that allows checking the integrity of network coded data. These works provide computational security (i.e., the attacker's resources are limited) in network coding.

Besides digital signatures and MACs, authentication codes also satisfy the properties of authentication. However, authentication code provides unconditional security (i.e., the attacker has unlimited computational power). In the multi-receiver authentication model, a sender broadcasts an authenticated message such that all the receivers can independently verify the authenticity of the message with their own private keys. It requires a security that malicious groups of up to a given size of receivers can not successfully impersonate the transmitter, or substitute a transmitted message. Desmedt *et al.* [14] gave an authentication scheme of single message for multi-receivers. Safavi-Naini and Wang [15] extended the DFY scheme [14] to be an authentication scheme of multiple messages for multi-receivers. Note that their construction was not linear over the base field with respect to the message. Oggier and Fathi [16, 1] made a little modification of the construction so that the construction can be used for network coding, which is actually not secure we will show in this paper. Tang [17] used homomorphic authentication codes to sign a subspace which provide an unconditionally security. In fact, Tang in the same paper had noticed that linear authentication codes for linear network is not secure, so he modified the type of substitution attack.

Firstly, we recall the general model of network coding and the definition of subspace codes. In the basic multicast model for linear network coding, a source node s generates n messages, each consisting of m symbols in the base field \mathbb{F}_q . Let $\{x_1, x_2, \dots, x_n\} \subseteq \mathbb{F}_q^{l \times 1}$ represent the set of messages. Based on the messages, the source node s transmits a message over each outgoing channel. At a node in the network, the symbols on its outgoing channel are \mathbb{F}_q -linear combinations of incoming symbols. For a node i , define $Out(i) = \{e \in E : e \text{ is an outgoing channel of } i\}$, and $In(i) = \{e \in E : e \text{ is an incoming channel of } i\}$. If the channel e of network carries packet $y(e)$, where $e \in Out(i)$, and i is an internal nodes, then $y(e)$ satisfies $y(e) = \sum_{d \in In(i)} k_{de} y(d)$. The $|In(i)| \times |Out(i)|$ matrix $K_i = [k_{de}]_{d \in In(i), e \in Out(i)}$ is called the *local encoding kernel at node i* . Note that each $y(e)$ is a linear combination of the messages sent by the source node, so there exists a vector

$f_e \in \mathbb{F}_q^{1 \times n}$ such that

$$y(e) = f_e \underline{\mathbf{X}}, \text{ where } \underline{\mathbf{X}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

The vector f_e is called the *global encoding vector* of channel e . Given the local encoding kernels for all the channels in network, the global encoding kernels can be calculated recursively in any upstream-to-downstream order as follows

$$f_e = \sum_{d \in \text{in}(i)} k_{de} f_d.$$

Write the received vectors at a node t as a column vector

$$A_t = (y(e) : e \in \text{In}(t))^T = \begin{pmatrix} y(e_1) \\ y(e_2) \\ \vdots \\ y(e_{e(t)}) \end{pmatrix},$$

where $\text{In}(t) = \{e_1, e_2, \dots, e_{e(t)}\}$. Then we have the decoding equation at the node t

$$F_t \cdot \underline{\mathbf{X}} = A_t,$$

where

$$F_t = (f_e : e \in \text{In}(t))^T = \begin{pmatrix} f_{e_1} \\ f_{e_2} \\ \vdots \\ f_{e_{e(t)}} \end{pmatrix}$$

is called the *global encoding kernel* at the node t .

2. THE AUTHENTICATION SCHEME OF OGGIER AND FATHI

Oggier and Fathi constructed an authentication code against pollution and substitution attacks in network coding, and they proved that the scheme is unconditional secure under some condition. Let us recall their construction and their result about the security analysis.

- **Key generation:** A trusted authority randomly generates $M + 1$ polynomials $P_0(x), P_1(x), \dots, P_M(x) \in \mathbb{F}_{q^l}[x]$ and choose V distinct values $x_1, \dots, x_V \in \mathbb{F}_{q^l}$. These polynomials are of degree $k - 1$, and we denote them by

$$P_i(x) = a_{i,0} + a_{i,1}x + a_{i,2}x^2 + \dots + a_{i,k-1}x^{k-1}, \quad i = 0, 1, \dots, M.$$

- **Key distribution:** The trusted authority gives as private key to the source S the $M + 1$ polynomials $(P_0(x), \dots, P_M(x))$, and as private key for each verifier R_i the $M + 1$ valuations of polynomials at $x = x_i$, namely $(P_0(x_i), \dots, P_M(x_i))$, $i = 1, 2, \dots, V$. The values x_1, \dots, x_V are made public. The keys can be given to the nodes when they sign up for a service protected by this scheme.
- **Authentication tag:** Let us assume that the source wants to send n data messages $s_1, s_2, \dots, s_n \in \mathbb{F}_{q^l}$. Choose and fix an \mathbb{F}_q -linear isomorphism between $\mathbb{F}_{q^l}^l$ and \mathbb{F}_{q^l} , then consider they have the same elements. The source computes the following polynomial in $\mathbb{F}_{q^l}[x]$:

$$A_{s_i}(x) = P_0(x) + s_i P_1(x) + s_i^q P_2(x) + \dots + s_i^{q^{M-1}} P_M(x)$$

which forms the authentication tag of each s_i , $i = 1, \dots, n$. Instead of sending the original messages s_1, s_2, \dots, s_n , the source actually sends packets \vec{x}_i of the form

$$\vec{x}_i = [1, s_i, A_{s_i}(x)] \in \mathbb{F}_q^{1+l+kl}, \quad i = 1, \dots, n.$$

The security of the authentication scheme above proven by Oggier and Fathi is as follows:

Proposition 2.1 ([1]). *Consider a multicast network implementing linear network coding, among which nodes V of them are verifying nodes owning a private key for authentication. The above scheme is an unconditionally secure network coding authentication code against a coalition of up to $k - 1$ adversaries, possibly among the verifying nodes, in which every key can be used to authentication up to M messages, under the assumption that $H \leq M$, where H is the sum of numbers of the incoming edges at each adversary.*

3. LINEAR SUBSTITUTION/POLLUTION ATTACKS TO THEIR SCHEME

In the security analysis given by Oggier and Fathi, they focused on solving the system of linear equations on variables $a_{i,j}$ to recover the private key of other node. Actually, notice that the authenticated vectors \vec{x}_i above are nearly linear on messages, so we can implement linear substitution attack to their scheme. In some papers[??], they have noticed that it is not secure to use linear authentication codes on linear network. And they considered a new type of substitution attack. Also, they pointed out that the authentication code of Oggier and Fathi is non-linear so that it should be still secure. Next, we present our linear substitution attack in details.

Suppose the coalition of malicious verifying nodes can carry out decoding of the network coding, i.e., the coalition of their global kernels has rank not less than the minimum cut of the network, for instance, the coalition of malicious verifying nodes contains one destination node. In this case, they can decode the tagged messages sent by the source node:

$$\vec{x}_i = [1, s_i, A_{s_i}(x)] \quad \text{for } i = 1, 2, \dots, n.$$

For any $a_1, a_2, \dots, a_n \in \mathbb{F}_q$ such that

$$a_1 + a_2 + \dots + a_n = 1,$$

replace \vec{x}_n by $\vec{x}_n^\# = \sum_{i=1}^n a_i \vec{x}_i$. Next, we show that in this way each verifying node can not notice this substitution attack.

Verification of Linear Substitution Attack:

The vector of any incoming edge at any node is of the form

$$\sum_{i=1}^{n-1} \alpha_i \vec{x}_i + \alpha_n \vec{x}_n^\# = \left[\sum_{i=1}^n \alpha_i, \sum_{i=1}^{n-1} \alpha_i s_i + \alpha_n s'_n, \sum_{i=1}^{n-1} \alpha_i A_{s_i}(x) + \alpha_n A_{s'_n}(x) \right]$$

for some $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{F}_q$. Then

$$\begin{aligned} & \sum_{i=1}^M \left(\sum_{j=1}^{n-1} \alpha_j s_j + \alpha_n s'_n \right)^{q^{i-1}} P_i(x) + P_0(x) \left(\sum_{j=1}^n \alpha_j \right) \\ &= \sum_{i=1}^M \left(\sum_{j=1}^{n-1} \alpha_j s_j^{q^{i-1}} + \alpha_n s_n'^{q^{i-1}} \right) P_i(x) + P_0(x) \left(\sum_{j=1}^n \alpha_j \right) \\ &= \sum_{i=1}^M \left(\sum_{j=1}^{n-1} \alpha_j s_j^{q^{i-1}} + \alpha_n \sum_{t=1}^n a_t s_t^{q^{i-1}} \right) P_i(x) + P_0(x) \left(\sum_{j=1}^n \alpha_j \right) \\ &= \sum_{i=1}^M P_i(x) \left(\sum_{j=1}^{n-1} \alpha_j s_j^{q^{i-1}} \right) + \alpha_n \sum_{i=1}^M P_i(x) \left(\sum_{t=1}^n a_t s_t^{q^{i-1}} \right) \\ & \quad + P_0(x) \left(\sum_{j=1}^n \alpha_j \right) \end{aligned}$$

equals to

$$\begin{aligned}
& \sum_{i=1}^{n-1} \alpha_i A_{s_i}(x) + \alpha_n A_{s'_n}(x) \\
= & \sum_{i=1}^{n-1} \alpha_i \left(P_0(x) + \sum_{t=1}^M s_i^{q^{t-1}} P_t(x) \right) \\
& + \alpha_n \left(\left(\sum_{j=1}^{n-1} a_j \right) P_0(x) + \sum_{t=1}^M s_n^{q^{t-1}} P_t(x) \right) \\
= & \left(\sum_{i=1}^n \alpha_i \right) P_0(x) + \sum_{i=1}^{n-1} \sum_{t=1}^M \alpha_i s_i^{q^{t-1}} P_t(x) \\
& + \alpha_n \sum_{t=1}^M \left(\sum_{j=1}^n a_j s_j^{q^{t-1}} \right) P_t(x) \\
= & \left(\sum_{i=1}^n \alpha_i \right) P_0(x) + \sum_{t=1}^M \left(\sum_{i=1}^{n-1} \alpha_i s_i^{q^{t-1}} \right) P_t(x) \\
& + \alpha_n \sum_{t=1}^M \left(\sum_{j=1}^n a_j s_j^{q^{t-1}} \right) P_t(x)
\end{aligned}$$

for all $x \in \mathbb{F}_{q^t}$. In other words, it can be verified by any verifying node using his private key.

From the above argument, we can see that any node in the network can easily make pollution to the network flow in the way that the node replaces any one or more of the vectors he received by linear combinations of his incoming vectors whose coefficients have sum 1 and then the node processes the network coding with the new vectors.

Finally, we point out that even if Oggier and Fathi's scheme can work fruitfully, the condition $H \leq M$ can also be removed. Note that the condition $H \leq M$ is very critical in a network. The proof is similar to the proof given by Oggier and Fathi. They wrote the secret parameters $A = (a_{i,j})$ as a column vector in the order as following

$$\vec{a} = (a_{0,1}, a_{0,2}, \dots, a_{0,k}, a_{1,1}, \dots, a_{1,k}, \dots, a_{M,1}, a_{M,2}, \dots, a_{M,k})^T,$$

where G^T represents the transpose of the matrix G , and they rewrote the system of linear equations using \vec{a} . Then they computed the rank of the coefficient matrix, finally they concluded that under the condition $H \leq M$ the rank of the coefficient matrix is less than the number of variables $k(M+1)$. Actually, if we rewrite the secret parameters $A = (a_{i,j})$ as a column vector in the following order

$$\vec{a}' = (a_{0,1}, a_{1,1}, \dots, a_{M,1}, a_{0,2}, \dots, a_{M,2}, \dots, a_{0,k}, a_{1,k}, \dots, a_{M,k})^T.$$

Then we obtain a new system of linear equations on $a_{i,j}$ using \vec{a}' . In this way, we can easily show that the rank of the coefficient matrix is always less than the number of variables. So the system of linear equations does always have solutions. Next, we give the details.

Suppose a group of K malicious nodes collaborate to recover A and make a substitution attack. Without loss of generality, we assume that the malicious nodes are R_1, R_2, \dots, R_K . Suppose the global encoding kernel at the verifying node R_i is

$$H_i = \begin{pmatrix} h_{1,1}^{(i)} & h_{1,2}^{(i)} & \dots & h_{1,n}^{(i)} \\ h_{2,1}^{(i)} & h_{2,2}^{(i)} & \dots & h_{2,n}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ h_{e(i),1}^{(i)} & h_{e(i),2}^{(i)} & \dots & h_{e(i),n}^{(i)} \end{pmatrix}.$$

Each R_i has some information about the secret parameter matrix $A = (a_{i,j})$:

$$\begin{pmatrix} \sum_{j=1}^n h_{1,j}^{(i)} & \sum_{j=1}^n h_{1,j}^{(i)} s_j & \sum_{j=1}^n h_{1,j}^{(i)} s_j^q & \cdots & \sum_{j=1}^n h_{1,j}^{(i)} s_j^{q^{M-1}} \\ \sum_{j=1}^n h_{2,j}^{(i)} & \sum_{j=1}^n h_{2,j}^{(i)} s_j & \sum_{j=1}^n h_{2,j}^{(i)} s_j^q & \cdots & \sum_{j=1}^n h_{2,j}^{(i)} s_j^{q^{M-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^n h_{e(i),j}^{(i)} & \sum_{j=1}^n h_{e(i),j}^{(i)} s_j & \sum_{j=1}^n h_{e(i),j}^{(i)} s_j^q & \cdots & \sum_{j=1}^n h_{e(i),j}^{(i)} s_j^{q^{M-1}} \end{pmatrix} \cdot A \\ = \begin{pmatrix} \sum_{j=1}^n h_{1,j}^{(i)} L_1(s_j) & \sum_{j=1}^n h_{1,j}^{(i)} L_2(s_j) & \cdots & \sum_{j=1}^n h_{1,j}^{(i)} L_k(s_j) \\ \sum_{j=1}^n h_{2,j}^{(i)} L_1(s_j) & \sum_{j=1}^n h_{2,j}^{(i)} L_2(s_j) & \cdots & \sum_{j=1}^n h_{2,j}^{(i)} L_k(s_j) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^n h_{e(i),j}^{(i)} L_1(s_j) & \sum_{j=1}^n h_{e(i),j}^{(i)} L_2(s_j) & \cdots & \sum_{j=1}^n h_{e(i),j}^{(i)} L_k(s_j) \end{pmatrix}$$

and

$$A \cdot \begin{pmatrix} 1 \\ x_i \\ \vdots \\ x_i^{k-1} \end{pmatrix} = \begin{pmatrix} P_0(x_i) \\ P_1(x_i) \\ \vdots \\ P_M(x_i) \end{pmatrix}.$$

The group of malicious nodes combines their equations, and they get a system of linear equations

$$(1) \quad \begin{cases} \begin{pmatrix} D_1 \\ \vdots \\ D_K \end{pmatrix} \cdot A = \begin{pmatrix} C_1 \\ \vdots \\ C_K \end{pmatrix}, \\ A \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1^q & x_2^q & \cdots & x_K^q \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{q^{k-1}} & x_2^{q^{k-1}} & \cdots & x_K^{q^{k-1}} \end{pmatrix} = \begin{pmatrix} P_0(x_1) & P_0(x_2) & \cdots & P_0(x_K) \\ P_1(x_1) & P_1(x_2) & \cdots & P_1(x_K) \\ \vdots & \vdots & \ddots & \vdots \\ P_M(x_1) & P_M(x_2) & \cdots & P_M(x_K) \end{pmatrix}, \end{cases}$$

where

$$D_i = \begin{pmatrix} \sum_{j=1}^n h_{1,j}^{(i)} & \sum_{j=1}^n h_{1,j}^{(i)} s_j & \sum_{j=1}^n h_{1,j}^{(i)} s_j^q & \cdots & \sum_{j=1}^n h_{1,j}^{(i)} s_j^{q^{M-1}} \\ \sum_{j=1}^n h_{2,j}^{(i)} & \sum_{j=1}^n h_{2,j}^{(i)} s_j & \sum_{j=1}^n h_{2,j}^{(i)} s_j^q & \cdots & \sum_{j=1}^n h_{2,j}^{(i)} s_j^{q^{M-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^n h_{e(i),j}^{(i)} & \sum_{j=1}^n h_{e(i),j}^{(i)} s_j & \sum_{j=1}^n h_{e(i),j}^{(i)} s_j^q & \cdots & \sum_{j=1}^n h_{e(i),j}^{(i)} s_j^{q^{M-1}} \end{pmatrix}$$

and

$$C_i = \begin{pmatrix} \sum_{j=1}^n h_{1,j}^{(i)} L_1(s_j) & \sum_{j=1}^n h_{1,j}^{(i)} L_2(s_j) & \cdots & \sum_{j=1}^n h_{1,j}^{(i)} L_k(s_j) \\ \sum_{j=1}^n h_{2,j}^{(i)} L_1(s_j) & \sum_{j=1}^n h_{2,j}^{(i)} L_2(s_j) & \cdots & \sum_{j=1}^n h_{2,j}^{(i)} L_k(s_j) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^n h_{e(i),j}^{(i)} L_1(s_j) & \sum_{j=1}^n h_{e(i),j}^{(i)} L_2(s_j) & \cdots & \sum_{j=1}^n h_{e(i),j}^{(i)} L_k(s_j) \end{pmatrix}.$$

Denote

$$S_n = \begin{pmatrix} 1 & s_1 & s_1^q & \cdots & s_1^{q^{M-1}} \\ 1 & s_2 & s_2^q & \cdots & s_2^{q^{M-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & s_n & s_n^q & \cdots & s_n^{q^{M-1}} \end{pmatrix}.$$

Then

$$D_i = H_i \cdot S_n.$$

Lemma 3.1. *If $K \leq k - 1$, then there exists exact $q^{l(M+1-r_0)(k-K)}$ matrices A satisfying the system of equations (1), where*

$$r_0 = \text{rank} \left(\begin{pmatrix} H_1 S_n \\ H_2 S_n \\ \vdots \\ H_K S_n \end{pmatrix} \right).$$

Proof. Recall the system (1)

$$\left\{ \begin{array}{l} \begin{pmatrix} H_1 S_n \\ \vdots \\ H_K S_n \end{pmatrix} \cdot A = \begin{pmatrix} C_1 \\ \vdots \\ C_K \end{pmatrix}, \\ A \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1^q & x_2^q & \cdots & x_K^q \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{q^{k-1}} & x_2^{q^{k-1}} & \cdots & x_K^{q^{k-1}} \end{pmatrix} = \begin{pmatrix} P_0(x_1) & \cdots & P_0(x_K) \\ P_1(x_1) & \cdots & P_1(x_K) \\ \vdots & \ddots & \vdots \\ P_M(x_1) & \cdots & P_M(x_K) \end{pmatrix} \end{array} \right.$$

Rewrite the matrix A of variables as a single column of $k(M + 1)$ variables. Then the system (1) becomes

$$(2) \quad \begin{pmatrix} H_1 S_n & 0 & 0 & 0 \\ 0 & H_1 S_n & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & H_1 S_n \\ \vdots & \vdots & \ddots & \vdots \\ H_K S_n & 0 & 0 & 0 \\ 0 & H_K S_n & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & H_K S_n \\ I_{M+1} & x_1 I_{M+1} & \cdots & x_1^{k-1} I_{M+1} \\ I_{M+1} & x_2 I_{M+1} & \cdots & x_2^{k-1} I_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ I_{M+1} & x_K I_{M+1} & \cdots & x_K^{k-1} I_{M+1} \end{pmatrix} \cdot \begin{pmatrix} a_{0,1} \\ a_{1,1} \\ \vdots \\ a_{M,1} \\ a_{0,2} \\ a_{1,2} \\ \vdots \\ a_{M,2} \\ \vdots \\ a_{0,k} \\ a_{1,k} \\ \vdots \\ a_{M,k} \end{pmatrix} = T$$

where I_{M+1} is the identity matrix with rank $(M + 1)$ and T is the column vector of the constant terms in system (1) with proper order. Notice that

$$r_0 = \text{rank} \left(\begin{pmatrix} H_1 S_n \\ H_2 S_n \\ \vdots \\ H_K S_n \end{pmatrix} \right) = \text{rank} \left(\begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_K \end{pmatrix} \cdot S_n \right) \leq \min \left\{ \text{rank} \left(\begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_K \end{pmatrix} \right), n \right\}.$$

Also note that rows of

$$\begin{pmatrix} H_1 S_n \\ H_2 S_n \\ \vdots \\ H_K S_n \end{pmatrix}$$

is contained in the space \mathbb{F}_q^{M+1} generated by $x_i^j I_{M+1}$ if $x_i \neq 0$. So the rank of the coefficient matrix of System (2) to

$$r_0 k + (M + 1 - r_0)K$$

which is less than the number of variables $k(M + 1)$. So the system (2) has

$$q^{l(k(M+1)-(r_0 k+(M+1-r_0)K))} = q^{l(M+1-r_0)(k-K)}$$

solutions, i.e., the system (1) has $q^{l(M+1-r_0)(k-K)}$ solutions. \square

4. CONCLUSION

In this paper, we discuss the security of the authentication code given by Oggier and Fathi and show our linear attack to their scheme, although it looks like non-linear. So we point out that as the technique of linear network develops very fast, and it has invaded a lot in our daily life, such as Internet TV, wireless networks, content distribution networks, P2P networks and distributed file system, to give an efficient and unconditional secure authentication code for linear network against the original substitution/pollution attack considered by Oggier and Fathi is extremely urgent.

REFERENCES

- [1] F. E. Oggier and H. Fathi, "An authentication code against pollution attacks in network coding," *IEEE/ACM Trans. Netw.*, vol. 19, no. 6, pp. 1587–1596, 2011.
- [2] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *Information Theory, IEEE Transactions on*, vol. 46, no. 4, pp. 1204–1216, jul 2000.
- [3] S.-Y. Li, R. Yeung, and N. Cai, "Linear network coding," *Information Theory, IEEE Transactions on*, vol. 49, no. 2, pp. 371–381, feb. 2003.
- [4] T. Ho, R. Koetter, M. Medard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *Information Theory, 2003. Proceedings. IEEE International Symposium on*, june-4 july 2003, p. 442.
- [5] M. Krohn, M. Freedman, and D. Mazieres, "On-the-fly verification of rateless erasure codes for efficient content distribution," in *Security and Privacy, 2004. Proceedings. 2004 IEEE Symposium on*, may 2004, pp. 226–240.
- [6] C. Gkantsidis and P. Rodriguez Rodriguez, "Cooperative security for network coding file distribution," in *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, april 2006, pp. 1–13.
- [7] Z. Yu, Y. Wei, B. Ramkumar, and Y. Guan, "An efficient signature-based scheme for securing network coding against pollution attacks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, april 2008, pp. 1409–1417.
- [8] A. Yun, J. H. Cheon, and Y. Kim, "On homomorphic signatures for network coding," *Computers, IEEE Transactions on*, vol. 59, no. 9, pp. 1295–1296, sept. 2010.
- [9] D. Charles, K. Jain, and K. Lauter, "Signatures for network coding," in *Information Sciences and Systems, 2006 40th Annual Conference on*, march 2006, pp. 857–863.
- [10] F. Zhao, T. Kalker, M. Medard, and K. J. Han, "Signatures for content distribution with network coding," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, june 2007, pp. 556–560.
- [11] D. Boneh, D. Freeman, J. Katz, and B. Waters, "Signing a linear subspace: Signature schemes for network coding," in *Public Key Cryptography C PKC 2009*, ser. Lecture Notes in Computer Science, S. Jarecki and G. Tsudik, Eds. Springer Berlin / Heidelberg, 2009, vol. 5443, pp. 68–87. [Online]. Available: <http://dx.doi.org/10.1007/978-3-642-00468-1-5>

- [12] D. Boneh and D. Freeman, “Linearly homomorphic signatures over binary fields and new tools for lattice-based signatures,” in *Public Key Cryptography C PKC 2011*, ser. Lecture Notes in Computer Science, D. Catalano, N. Fazio, R. Gennaro, and A. Nicolosi, Eds. Springer Berlin / Heidelberg, 2011, vol. 6571, pp. 1–16. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-19379-8_1
- [13] S. Agrawal and D. Boneh, “Homomorphic MACs: MAC-based integrity for network coding,” in *Applied Cryptography and Network Security*, ser. Lecture Notes in Computer Science, M. Abdalla, D. Pointcheval, P.-A. Fouque, and D. Vergnaud, Eds. Springer Berlin / Heidelberg, 2009, vol. 5536, pp. 292–305. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-01957-9_18
- [14] Y. Desmedt, Y. Frankel, and M. Yung, “Multi-receiver/multi-sender network security: efficient authenticated multicast/feedback,” in *INFOCOM '92. Eleventh Annual Joint Conference of the IEEE Computer and Communications Societies, IEEE*, may 1992, pp. 2045–2054 vol.3.
- [15] R. Safavi-naini and H. Wang, “New results on multi-receiver authentication codes,” in *Advances in Cryptology – EUROCRYPT '98, LNCS*. Springer-Verlag, 1998, pp. 527–541.
- [16] F. Oggier and H. Fathi, “Multi-receiver authentication code for network coding,” in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, sept. 2008, pp. 1225–1231.
- [17] Z. Tang, “Homomorphic A-codes for network coding,” Cryptology ePrint Archive, Report 2012/331, 2012, <http://eprint.iacr.org/>.

CHERN INSTITUTE OF MATHEMATICS, NANKAI UNIVERSITY, TIANJIN, P.R. CHINA

E-mail address: zhangjun04@mail.nankai.edu.cn; xinranli@mail.nankai.edu.cn; fwfu@nankai.edu.cn